REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-00-

0445

Public reporting burden for this collection of information is estimated to average 1 hour per response, i... gathering and maintaining the data needed, and completing and reviewing the collection of informatio... collection of information, including suggestions for reducing this burden, to Washington Headquarters... Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, P... ...ta sources, ...pect of this ...5 Jefferson

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE 9/5/00	3. REPORT TYPE AND DATES COVERED Final 5/1/96 - 9/30/98

4. TITLE AND SUBTITLE
Further Research on Super Auditory Localization for Improved Human-Machine Interfaces

5. FUNDING NUMBERS
F49620-96-1-0202

6. AUTHOR(S)
Nathaniel Durlach

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Research Laboratory of Electronics
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research

801 North Randolph St, Rm 732
Arlington, VA 22203-1977

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

We achieved goals 1-6, 8 and 11. Due to the nature of our findings, we did not pursue goals 7, 9 and 10; instead, we focused on the rich problem of near-source HRTFs and perception of nearby sources (area 11). In addition to measuring and anlayzing these HRTFs, we performed a series of experiments investigating localization of nearby sources.

20000922 050

14. SUBJECT TERMS
Human-machine interfaces, auditory

15. NUMBER OF PAGES
17

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

Further Research on Super Auditory Localization for Improved Human-Machine Interfaces

Final Report

5/1/96-9/30/98

Principal Investigator: Nathaniel Durlach
Massachusetts Institute of Technology
36-709
77 Massachusetts Ave.
Cambridge, MA 02139

Further Research on Super Auditory Localization for Improved Human-Machine Interfaces

Objectives

The general goals of the current project are (1) to determine, understand, and model the perceptual effects of altered auditory localization cues and (2) to design, construct, and evaluate cue alterations that can be used to improve performance of human-machine interfaces in virtual-environment and teleoperation systems. To the extent that the research is successful, it will both advance our understanding of auditory localization and adaptation, and improve our ability to design human-machine interfaces that provide a high level of performance.

The specific goals for the project are shown below.

1. Continue to acquire, develop, and integrate devices and facilities into our laboratory that are suitable for studying supernormal auditory localization.

2. Analyze in more detail the data already obtained using the azimuthal remapping transformation

$$\theta' = f_n(\theta) = \frac{1}{2}\tan^{-1}\left[\frac{2n\sin(2\theta)}{1-n^2+\left(1+n^2\right)\cos(2\theta)}\right]$$

with n = 3.

3. Conduct additional experiments using the $f_3(\theta)$ transformation to clarify and broaden the results already obtained with this transformation, focusing primarily on how visual cues affect adaptation.

4. Perform a series of experiments using the family of transformations $\{f_n(\theta)\}$ n = 1, 2, 3, 4, to study the effects of transformation severity and incremental exposure, as well as to explore questions related to conditional or dual adaptation.

5. Conduct a study similar to that mentioned in item 3 using the frequency-scaling family of transformations (which simulate increased head size) rather than the $f_n(\theta)$ family.

6. Refine our quantitative model of how resolution and bias in the localization of azimuth are influenced by transformations of azimuthal localization cues.

7. Conduct a further series of experiments, again using the frequency-scaling transformations, to determine the effects of such transformations on the perception of elevation.

8. Determine the potential of varying sets of filter transfer characteristics for coding distance and/or elevation by studying the information transfer that can be achieved using these sets.

9. Determine how the transformations that appear most promising in the previous studies, all of which evaluate the transformation in acoustic environments containing only a single target, perform in multiple-simultaneous-target environments.

10. Using one of the more promising sets of filter transfer characteristics studied in project 8, evaluate the usefulness of this set for coding distance and/or elevation by conducting adaptation experiments similar to those pursued in projects 3, 4, and 5 on the azimuthal variable.

11. Extend our knowledge of natural localization cues by measuring and analyzing Head-Related Transfer Functions (HRTFs) for sources in the near field (i.e., at a distance of 1 meter or less from the center of the head).

Status of Effort

We achieved goals 1-6, 8, and 11. Due to the nature of our findings, we did not pursue goals 7, 9, and 10; instead, we focussed on the rich problem of near-source HRTFs and perception of

nearby sources (area 11). In addition to measuring and analyzing these HRTFs, we performed a series of experiments investigating localization of nearby sources.

Accomplishments/New Findings

Model Predictions

A detailed description of our model of adaptation to supernormal auditory localization cues can be found in the attached paper entitled "Adaptation to Supernormal Auditory Localization Cues: A Decision-Theory Model." The main features of the model are summarized here.

The model assumes that all responses are determined by the value of an internal decision variable which is stochastic. The mean of the decision variable is monotonically related to the azimuth normally associated with a given stimulus, while its variance depends upon the range of stimuli being attended by the subject at a given point in time. More specifically, the variance in the underlying decision space grows as the attended range of stimuli grows. Criteria are placed along the uni-dimensional decision axis to divide the axis into n contiguous regions, corresponding to the n possible reponses on the n-alternative, forced-choice identification tasks used in our study. Adaptation occurs as the criteria shift, allowing subjects to change their mean respose to specific stimuli as the locations of the response regions move.

The model assumes that placement of the decision criteria and effective range (determining underlying variability in the decision variable) are determined by the relationship between physical stimuli and mean response at any given point in time. These changes are examined fully in the attached paper entitled "Adapting to Supernormal Auditory Localization Cues: II. Changes in Mean Response.". In this paper, it was shown that mean response is linearly related to the azimuth normally associated with a particular physical stimulus at all times. Changes in mean respose occur as the slope relating these values exponentially approaches an asymptotic value. The rate of change (reflecting the rate of adaptation) is statistically independent of the experimental factors investigated to date, while the asymptote depends upon the strength of the transformation employed.

With these assumptions, our model is able to fit total sensitivity (sum of d' acrosss all adjacent pairs of stimuli in an experiment), bias, and resolution for all experiments performed to date. The three figures below show the effectiveness of the model in fitting these quantities. It should be pointed out that the model paramters were chosen to fit Δ' (total sensitivity). The resulting parameter values were then used to predict bias and resolution (with good results).

Figure 1 shows the actual results (error bars) and the model predictions (open circles) for total sensitivity as a function of run. For each experiment, the two free model parameters were fit by finding values which minimized the mean square error between predictions and total sensitivity for each subject. These parameter values were then averaged across subjects to yield the model parameters used in all predictions. The figure shows predictions and results for five experiments, which differed mainly in the strength of the transformation employed (n=2, 3, or 4) and the number of source positions used (either [-60, +60] deg or [-30, 30] degs). Details of the differences across experiments can be found in the attached paper entitled "Adapting to Supernormal Auditory Localization Cues: I. Bias and Resolution."

Predictions match overall magnitude of the results well, although there is a tendency to underestimate sensitivity in some experiments (Experiments III and V, panels and b) and to overestimate sensitivity in Experiment VI (panel d). Identical model parameters were used to fit data from all experiments; thus, the fit across all experiments should be considered. The model also predicts abrupt changes in total sensitivity quite well (note results and predictions for Experiments IV and VII, panels b and e).

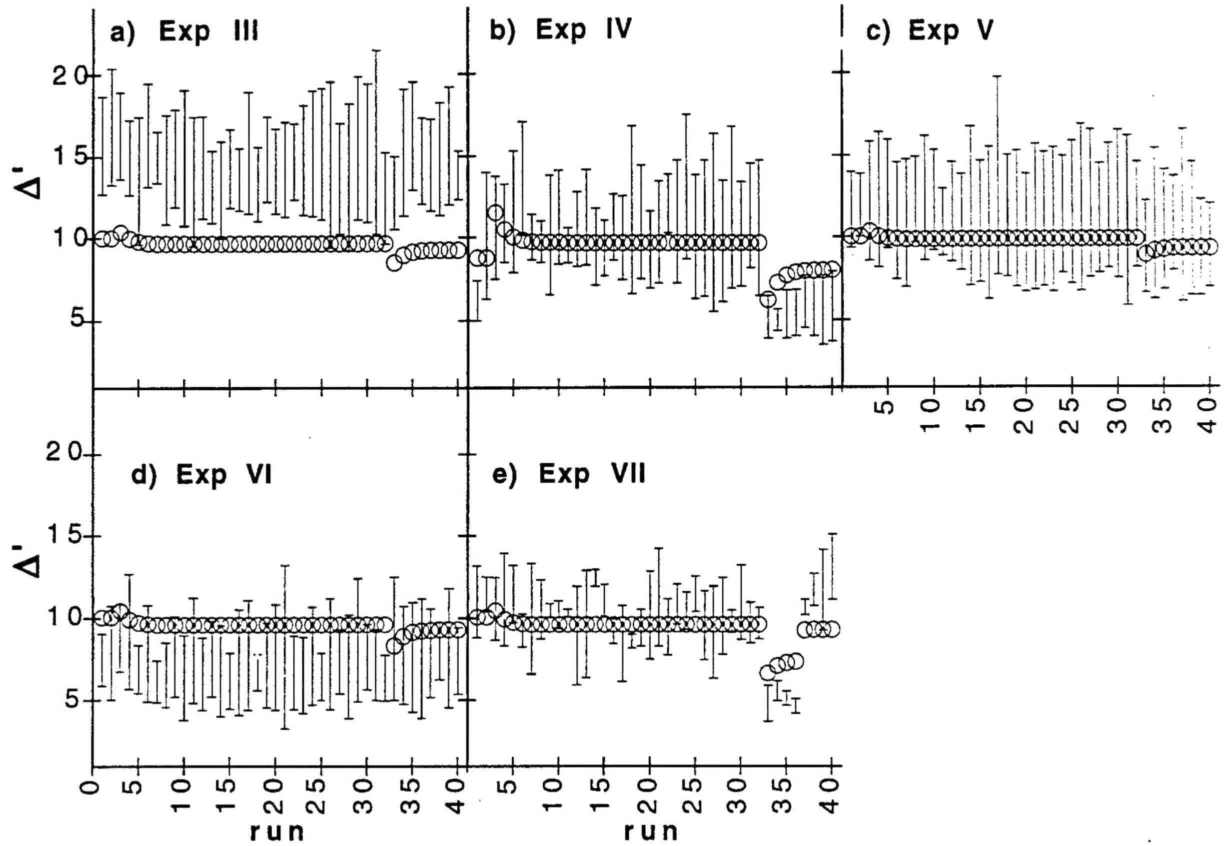

Figure 1. Predicted and actual Δ′ as a function of run (experimental data repeated from Figure 4). Open circles show predicted values, solid lines give mean experimental results (averaged across subjects) plus and minus one standard deviation.

Figure 2 shows the actual and predicted results for bias as a function of source position for the same experiments. In this figure and in Figure 3, open symbols show data prior to adaptation and filled symbols show results after adaptation. Circles represent normal-cue results and squares represent altered-cue results. In Figure 2, actual results are shown in the left panels and the corresponding model predictions shown in the right panels. The degree to which the left and right panels are alike is a measure of the degree to which the model predicts bias correctly.

In all experiments, a large bias is introduced when altered cues are first presented (open squares). This bias is reduced as subjects adapt (compare filled to open squares). At the end of the adaptation period, a large bias is seen which is opposite in sign to the bias first introduced with the altered cues (compare filled circles to open squares). Model predictions show the same trends for all experiments.

Finally, Figure 3 shows results and predictions for resolution as a function of source position. In Figure 3, the left side of each panel shows experimental results (averaged assuming data are left-right mirror symmetric) and the right side of each panel gives predictions from the model for positions right of center. The ability of the model to predict these results is measured by the degree to which each panel is mirror symmetric. Model parameters were identical to the values used to predict total sensitivity (Figure 1) and bias (Figure 2).

Again, the model predicts results quite closely. The use of supernormal cues produces an increase in resolution for positions near zero degrees azimuth (open squares). As subjects adapt to the supernormal cues, however, there is a slight tendency for resolution to decrease (compare filled

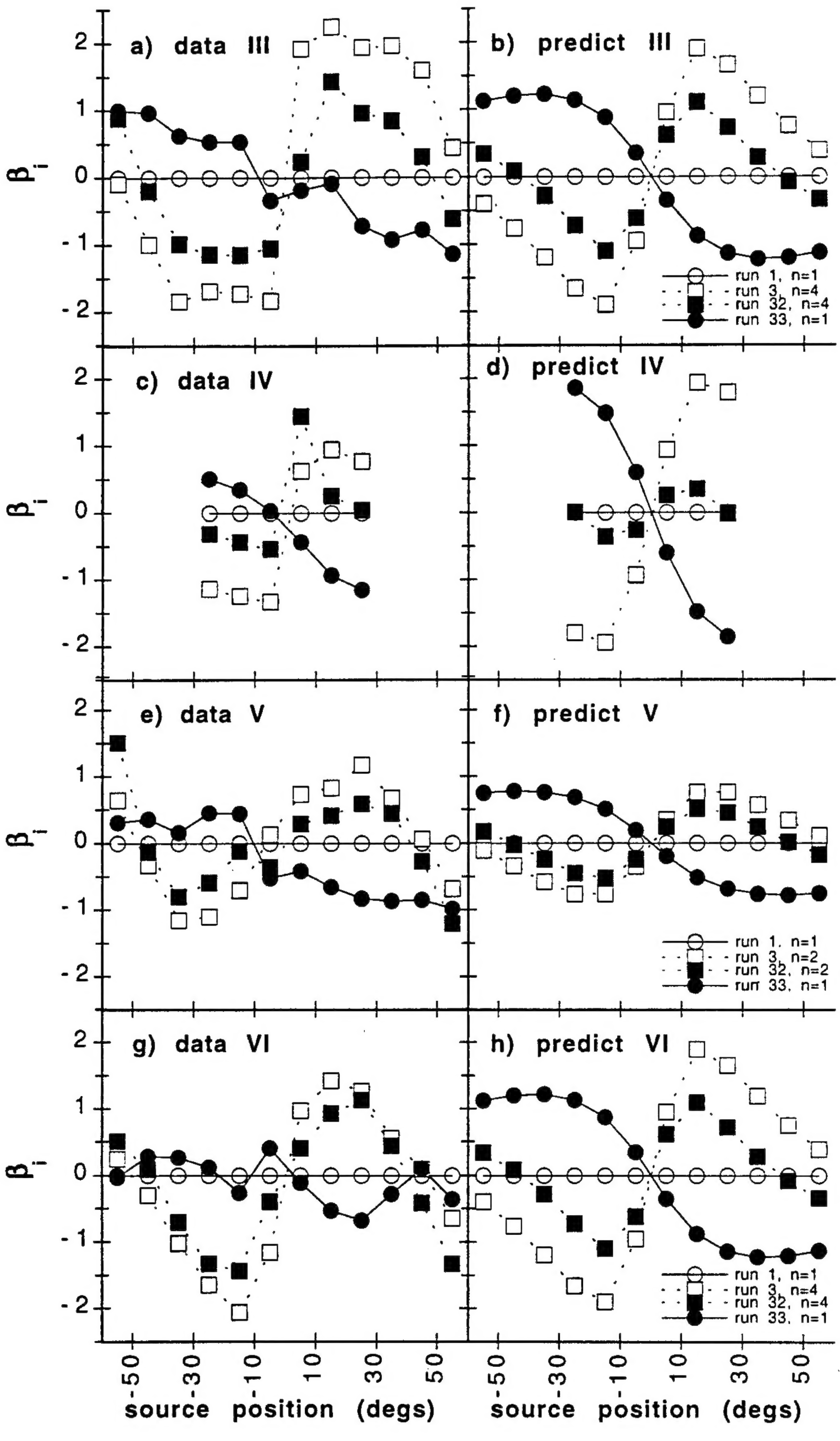

Figure 2. Actual and predicted bias as a function of position. Left panels show experimental results, right panels show corresponding model predictions. Open symbols represent results prior to training; filled symbols results after supernormal exposure. Circles represent normal-cue results; squares show altered-cue results.

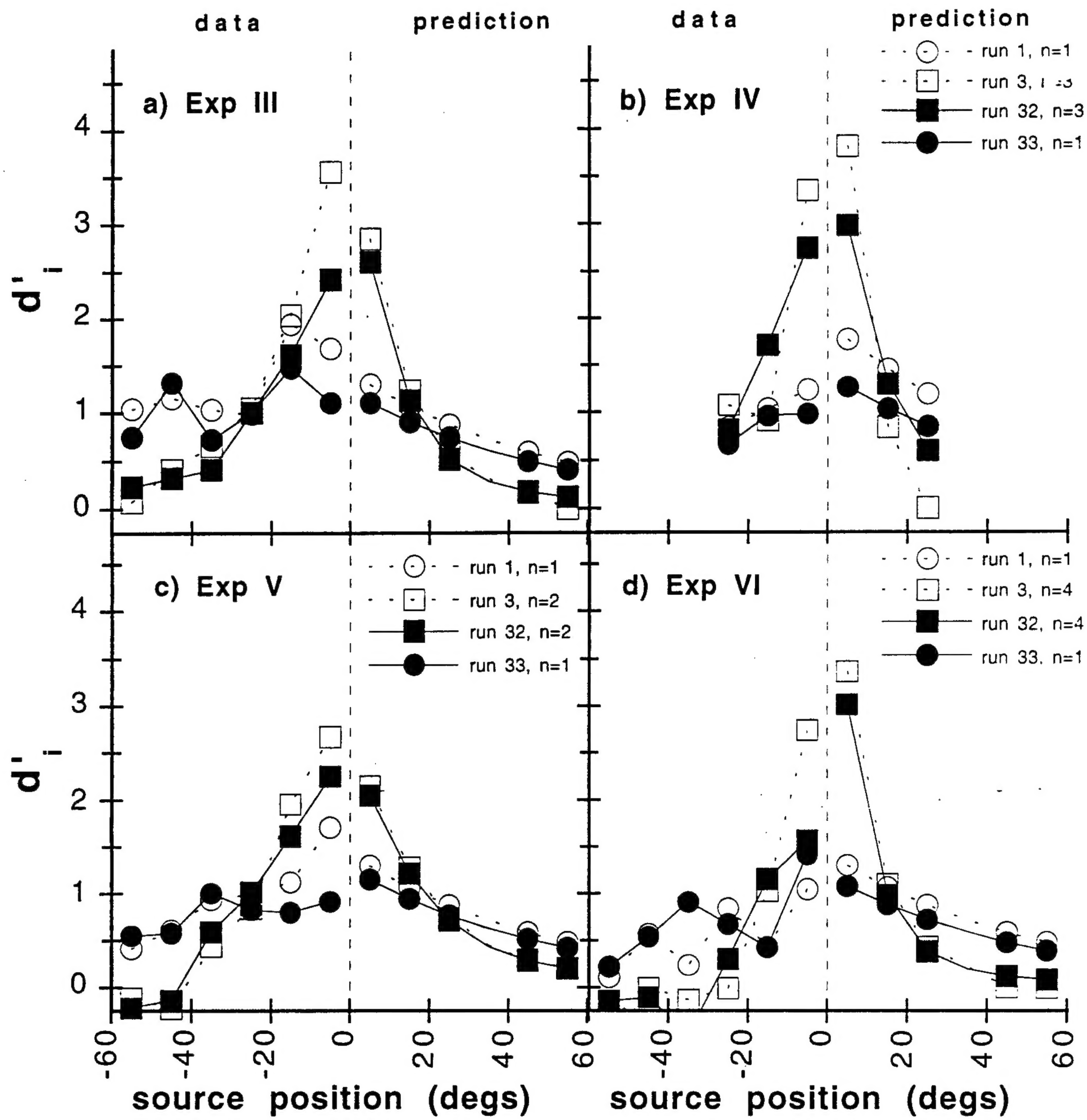

Figure 3. Actual and predicted d'_i as a function of position. Left side of each panel shows experimental results (averaged assuming data are left-right mirror symmetric). Right side of each panel gives predictions from the model for positions right of center. The dotted line separates the experimental data from the predicted results. Closeness of predictions is measured by the degree to which each panel is mirror symmetric. Again, the model predicts results quite closely. Model parameters were identical to the values used to predict total sensitivity (Figure 1) and bias (Figure 2).

to open squares). After adaptation, resolution using normal localization cues also tends to be worse than resolution prior to training (compare filled to open circles). The model predicts all of these trends. In addition, quantitative differences between experiments are also predicted by the model.

A more detailed discussion of both bias and resolution results can be found in the attached paper entitled "Adapting to Supernormal Auditory Localization Cues: I. Bias and Resolution." A

more complete discussion of the model and its ability to predict these results is given in "Adapting to Supernormal Auditory Localization : A Decision-Theory Model."

Effect of Visual Cues

Two previous experiments showed that adaptation to auditory-cue transformations depends upon visual information. In one experiment, subjects were blindfolded throughout the experimental session. When blindfolded, subjects could theoretically obtain information about the auditory-cue transformation by comparing the felt position of the head with auditory localization information. In the control experiment, subjects had a full visual field available to them (which could improve the accuracy with which the head position was registered) and were provided with explicit visual information (via a small lightbulb) as to the azimuthal position of the heard auditory source. In the blindfolded experiment, subjects showed no adaptation: localization errors were unchanged after 40 minutes of exposure to supernormal cues. In the control experiment, adaptation was found in all subjects. Two additional experiments have been performed to test exactly what type of visual information is necessary to allow subjects to adapt to transformed auditory localization cues.

The first experiment was identical to the previous blindfolded experiment, except that subjects viewed a normal visual field throughout the experiment, and the possible locations of the auditory source were marked by 13 light at the 13 possible source positions. Unlike the previous control experiment, subjects were not given any explicit visual information about the actual location of the heard auditory source (since the lights were never turned on). Preliminary analysis shows that subjects in this experiment adapted as fully as did subjects who were presented with explicit visual information about the location of the heard auditory source.

In the second experiment, the light arc (denoting the possible locations of the auditory sources) was removed. Thus, in these experiments, subjects had a complex visual field available to them at all times, but had no visual information about either the correct auditory source location, or the possible locations of the auditory sources. Preliminary analysis of these results indicate that subjects adapted in this condition, but that the adaptation was less complete than in the condition when the light arc marked all possible source locations.

Further analysis is necessary to fully quantify differences among these experiments; however, taken as a whole, these results indicate that the felt position of the head is not a sufficient cue to allow subjects to learn auditory cue rearrangements (cf., results from blindfolded experiment). However, when a visual field is present, subjects may be able to accurately register the location of their heads and can thus deduce how auditory localization cues change with changes in head position (cf. results from two new experiments). Finally, visual information about the possible locations of auditory sources also helps subjects to learn auditory cue remappings (compare results of two new experiments). Any additional benefit provided by explicit visual information as to the *correct* location of the auditory source is too small to be measured.

Effect of Changing Auditory-Cue Transformation

Two additional experiments have been completed which investigate what occurs when the strength of the transformation is changed half-way through the training period. In one experiment, a transformation of n=2 was used for the first 15 altered-cue tests and a transformation of n=4 was used for the final 15 altered-cue tests. In the second experiment, the order of the two transformations was reversed, with an n=4 transformation presented first and an n=2 transformation presented at the end of the training period.

Preliminary analysis suggests that, similar to results for single-transformation experiments, adaptation to multiple transformations can be measured by measuring the slope relating mean perceived location and physical cue. As in the single-transformation experiments, the slope appears to change exponentially towards an asymptotic value that depends only on the strength of the current transformation. Thus, in the n=2, n=4 experiment, subjects show an exponential decrease in slope towards the asymptote for n=2, then show a further exponential decrease in slope towards the asymptote for the transformation n=4. In the n=4, n=2 experiment, subjects show an exponential decrease towards the asymptote for n=4, then show an exponential increase in slope towards the asymptote for n=2. These changes in mean response can be fit by simple extension of the

exponential curve-fitting performed for single-transformation experiments (see the accompanying paper entitled "Adapting to Supernormal Localization Cues: II. Changes in Mean Response"). Further analysis is necessary to test whether bias and resolution results from these experiments can be predicted by our model of adaptation.

Near-Field HRTFs

In order to extend our knowledge of natural localization cues, we have begun to measure (and to test the use of) HRTFs for sources in the near field (i.e., at a distance of 1 meter or less). The localization cues that are available in this "near-field" region are different than those in the far field and have never been carefully studied. The results of this work are expected not only to increase our knowledge about the perception of direction for sources in the near field, but also to provide us with important information concerning the perception of distance (both for natural cue situations and for the development of supernormal distance cues).

Our study of near-field auditory localization cues has progressed in two major areas. A KEMAR acoustic manikin has been used to collect head-related transfer functions (HRTFs) for distances ranging from 0.15 m to 1 m from the center of the head, in cooperation with the Bioacoustics and Biocommunication branch of the Crew Systems Directorate of the Armstrong Laboratory (AL/CFBA) at Wright Patterson AFB. Measurements include azimuths and elevations from -45 deg to 45 deg. The most striking difference between these HRTFs and far-field HRTFs is an increase in interaural intensity differences (IIDs) as the sound source approaches the head. Futher work will investigate whether these distance-dependent IIDs may allow listeners to determine source distance in the near-field without reliance on overall intensity cues.

A pilot experiment designed to evaluate possible response methods for near-field auditory localization experiments was completed. Four response methods were compared: reporting coordinates verbally, pointing directly to the perceived location with a sensor on a wand, pointing to the perceived location relative to the location of a full-sized manikin head, and pointing to the perceived location relative to the location of a half-sized manikin head. Initial data analysis indicates that performance is best in the simple pointing task and worst when subjects must point to the location relative to the full-size manikin head.

Dual-State Adaptation

In previous grant years, we developed a quantitative model of adaptation that exhibits the following special features: (1) it deals with both resolution and response bias in a coherent and unified manner; (2) it includes consideration of the temporal course of both resolution and response bias as exposure time is increased and (3) it explains the generally-observed phenomenon of incomplete adaptation in terms of an inability to adapt to the non-linear components of a transformation (or, stated differently, adaptation is found to be complete to the hypothetical transformation that represents the best linear approximation to the given transformation).

In the past year, this preliminary model of adaptation has been further tested by examining the degree to which it can explain the performance of subjects who are asked to adapt to two different non-linear transformations during a single experimental session. In these experiments, subjects were presented with transformations from the $f_n(\theta)$ family identical to those that were used in previous experiments. In these dual-adaptation experiments, two experiments were performed. In the first experiment, subjects were presented with normal cues ($n = 1$), then a strong transformation ($n = 4$), then a weaker transformation ($n = 2$), and, finally, normal cues again ($n = 1$). In the second experiment, a different set of subjects heard normal cues ($n = 1$), then a weak transformation ($n = 2$), then a stronger transformation ($n = 4$), and, finally, normal cues again ($n = 1$). Results from these experiments were compared directly with results from experiments in which only the strong ($n = 4$) or weak ($n = 2$) transformation was employed, but which were otherwise identical to current experiments.

In our previous studies, we showed that the amount of adaptation achieved by the subject is summarized by a single "slope" parameter relating the normal position of a stimulus to currently-perceived position. This slope can be used to predict where a sound source is heard simply by multiplying the slope times the azimuthal position (angle) normally associated with the stimulus.

For example, normal perception corresponds to a slope of one. During adaptation to our transformations, subjects may achieve a slope of 0.8. When this occurs, a source normally associated with a position of 20 degrees will be heard (on average) near a location of 16 degrees (0.8 * 20 deg). In the single-transformation experiments, this slope parameter changed exponentially from near one (prior to adaptation) to an asymptotic value roughly equal to the best fit achievable for the given nonlinear transformation when it is assumed that subjects can only change their internal "slope" parameter.

In our year-one progress report, we indicated that, when the strength of the transformation is changed half-way through the training period, adaptation results are similar to what is expected based on results for single-transformation experiments. However, subsequent analysis has shown that there is a significant difference between what occurs in the dual-adaptation experiments and what might be expected from a simple extension of our model.

Figure 4 shows the internal slope parameter plotted as a function of run for four experiments using the transformations (a) n = 2, (b) n = 4, (c) n = 2 followed by n = 4, and (d) n = 4 followed by n = 2. For each individual subject, data across 8 identical sessions (gathered on 8 different days) was combined in order to have a large enough pool of data for analysis. These individual estimates were then averaged to yield the across-subject average, shown by the solid black line.

In our dual-adaptation experiments, adaptation is still summarized by a change in the single internal slope parameter. In addition, the time-course of changes in performance are consistent with the changes observed in the earlier experiments; that is, whenever the transformation of cues changes, subject performance changes exponentially with time to an asymptotic value. However, the asymptote found in the dual-adaptation experiments is not equal to the best-fit slope for the imposed transformation, as might be expected from a simple extension of the preliminary adaptation model previously developed. Instead, the asymptote appears to depend upon the experiment as well as on the transformation employed at a given time.

The most striking feature of these results is the demonstration of learning across sessions. On average, the asymptote reached by subjects during the *first* transformation period in the dual-state experiments is different from that shown by the population of comparable single-transformation subjects. More specifically, comparing results in panel c) with those in panel a), we see that subjects in the dual-state experiment appear to *overadapt* relative to both the single-state adaptation subjects and the best-fit slope for the n=2 transformation. Similarly, comparing results in panel d) with those in b), subjects in the dual-state adaptation experiment *overadapt* to the n=4 transformation compared to subjects in the single-state experiment and to the best-fit slope for the n=4 transformation. These differences occur even though the stimuli presented to subjects in the single- and dual-state experiments were identical through Run 17. Only in Run 18, when the second transformation was first presented to subjects in the dual-state experiments, did the two populations of subjects receive different stimuli.

Further analysis is necessary to clarify the ways in which learning carries over from day to day. This finding has important implications for understanding the rate of adaptation and degree of retention of learning, as well as how the introduction of multiple transformations affects adaptation. While we have always suspected that across-day learning occurs, we were unable to find such effects previously because these effects tend to be small relative to the within-day learning that occurs. Only by comparing results across experiments for different subject populations can we see the effect of learning across days. In future experiments, we hope to address this issue directly by having subjects perform two or three practice sessions (in which most across-day learning should occur) prior to gathering data in which we will examine within-day learning.

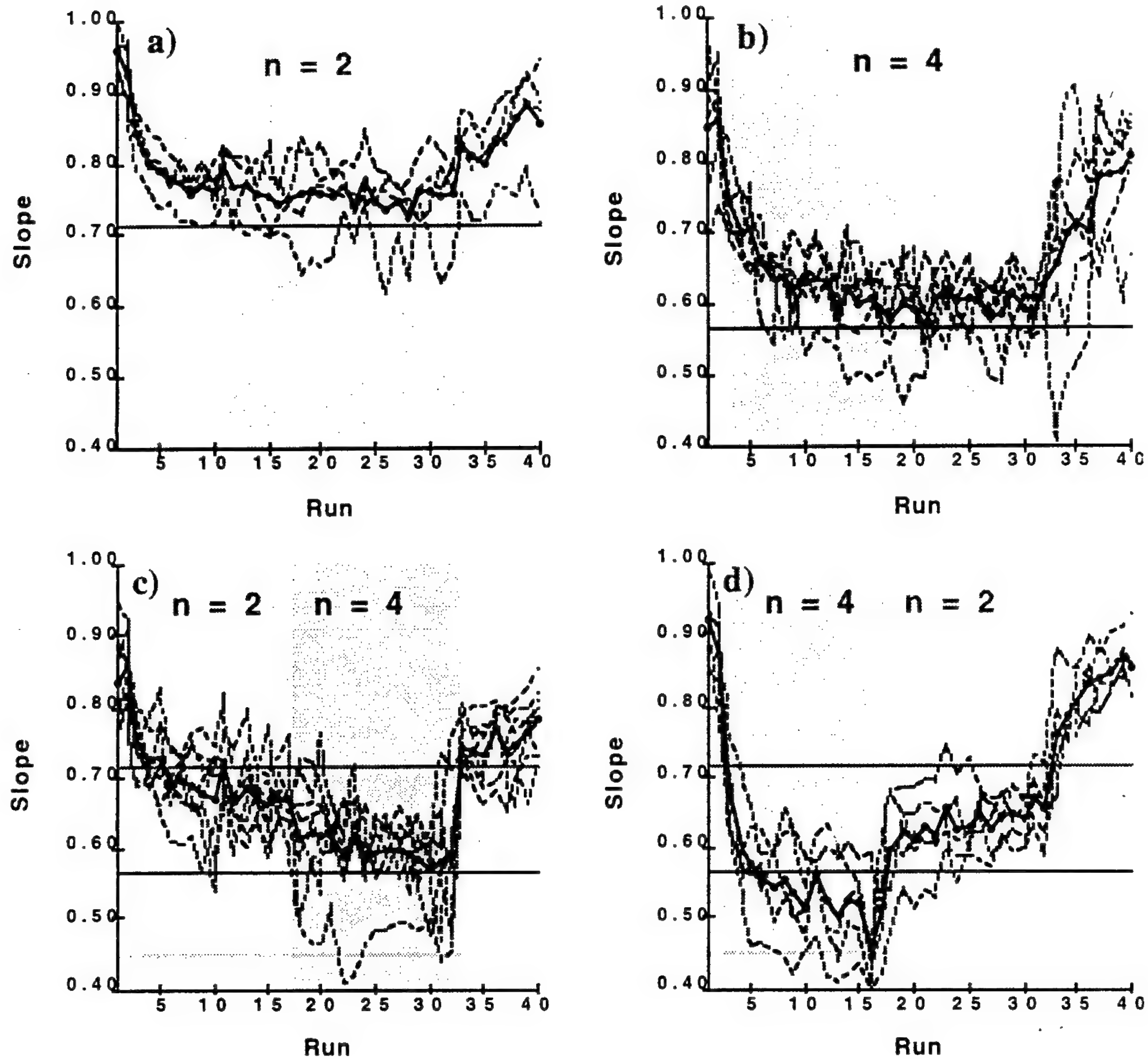

Figure 4. Individual subject and across-subject estimates of the best-fit slope as a function of run. Individual subjects given by dashed lines, across-subject averages by solid black line. Best-fit slope for the transformation(s) employed are shown by horizontal lines: pink is used to show the best-fit line for the n = 2 transformation, blue for the n = 4 transformation. The pink- and blue-shaded areas show which runs used the n = 2 and n = 4 transformation. respectively. a) n = 2 experiment, b) n = 4 experiment, c) n = 2 then n = 4 experiment, d) n = 4 then n = 2 experiment.

Adaptation to an Enlarged-Head Transformation

We have collected data from five subjects to investigate the degree to which subjects can adapt to HRTF cues like those that would result from a larger-than-normal head. The transformation employed in these experiments is given by the following equations:

$$H'_L(\omega,\theta,\phi) = H_L(K\omega,\theta,\phi)$$

$$H'_R(\omega,\theta,\phi) = H_R(K\omega,\theta,\phi)$$

where $H_L(\omega,\theta,\phi)$ and $H_R(\omega,\theta,\phi)$ are the normal far-field HRTFs, $H'_L(\omega,\theta,\phi)$ and $H'_R(\omega,\theta,\phi)$ are the transformed HRTFs, ω denotes frequency, θ denotes azimuth, ϕ denotes elevation, and K is a constant.

The extent to which the frequency scaling specified by the above equations (or, equivalently, the scaling of wavelengths) accurately simulates an enlarged head is discussed in detail in Rabinowitz, Maxwell, Shao and Wei (1993). In the next few months, we will analyze data from these "enlarged-head" adaptation experiments to determine how resolution and bias change over time as subjects adjust their performance. Particular attention will be paid to the degree to which subjects can extract information from localization cues that are not normally encountered in every day life (i.e., interaural time differences larger than those normally experienced).

Evaluation of Response Methods

Four response methods allowing subjects to indicate locations within one meter of a subject's head were evaluated experimentally. In the Direct-Location (DL) method, a subject moved a response pointer directly to the perceived target location. In the Large-Head (LH) method, the subject moved the response pointer to the perceived location relative to a manikin head corresponding to the location of the target relative to their own head. The Small-Head (SH) was similar to LH, except that a half-scale manikin head was used and the subjects were asked to scale down their responses by a factor of two. In the Verbal Report (VR) response, subjects verbally indicated the spherical coordinates of the target location. An experiment with a visual target indicated that the DL responses relatively unbiased and considerably more accurate than those of the other three methods. The three indirect methods, LH, SH, and VR, were all roughly equivalent in performance. Correcting for bias improved accuracy in the LH, SH, and VR responses, but not to the level of accuracy found in the DL responses.

When the visual target was replaced with an acoustic stimulus, the errors in the DL response were approximately doubled. In the acoustic experiment, the errors were approximately equivalent in the front and rear hemispheres, despite the expected difficulties of reaching behind the body and outside the visual field. The results suggest that the DL method is most appropriate for a near-field auditory localization experiment.

Near-Field Auditory Localization Experiments

We have completed data collection in a set of identification experiments designed to measure near-field auditory localization. In these experiments, which took place in MIT's anechoic chamber, an acoustic point source was placed at a random location by the experimenter prior to each trial. At the beginning of each trial, the location of the source was measured using an electromagnetic tracker, and the distance from the source to the listener's head was calculated. This distance was used to correct the amplitude of the stimulus signal (five 150 ms bursts of broadband noise) to make it independent of source distance. The stimulus amplitude was then randomized over an additional 15 dB range, and presented to the subject. After hearing the stimulus, the subject moved a response sensor to the perceived location of the stimulus, and the response location was then recorded by the control computer. Five conditions were tested:

1. Baseline: Standard experiment with broadband stimulus and random amplitude.
2. Monaural: An ear-plug and ear-muff occluded the contralateral ear.
3. Fixed Amplitude: The stimulus amplitude was fixed on each trial.
4. High-pass Filtered: The stimulus was high-pass filtered above 3 kHz.
5. Low-pass Filtered: The stimulus was low-pass filtered below 3 kHz.

Four subjects participated in the experiment. A total of 2000 trials were collected from each subject in the Baseline condition, and 500 trials were collected from each subject in the other conditions. The results for accuracy in direction in each condition are summarized in Table 1. Statistics are given for the signed errors in azimuth and elevation, the great arc angle from the stimulus position to the response position, and the percentage of trials resulting in front-back confusions. In each case the standard deviations were calculated separately for each subject before being averaged together to generate the data in the table. A front-back confusion was determined to

occur in a trial whenever the stimulus was in the front hemisphere and the response was in the rear hemisphere, or vise-versa. All reversals were corrected by reflecting the response across the frontal plane prior to calculation of the angular errors.

The Baseline data are roughly comparable to the localization identification experiment by Wightman and Kistler (1989) in the far field. They reported a mean angular error of 21° and 7% reversals. Our near-field experiments showed a slightly lower angle error and a higher reversal percentage than reported by Wightman and Kistler.

Error	*Baseline*	*Monaural*	*Fixed*	*Low-pass*	*High-pass*
Azimuth	1.0° (14.5)	-1.1° (31.2)	-1.8° (18.8)	-1.3° (17.8)	2.0° (19.4)
Elevation	2.5° (15.8)	0.0° (22.8)	-1.7° (13.9)	-10.9° (34.8)	3.1° (22.2)
Angle	17.1° (11.1)	29.9° (17.9)	17.2° (10.0)	33.6° (22.1)	23.5° (14.4)
Reversals	16%	22%	16%	43%	26%

Table 1: Mean Directional Errors in Preliminary Psychoacoustic Experiments (Standard Deviations are in Parentheses)

In the monaural condition, the standard deviation in azimuth and the angle error are large, but the standard deviation in elevation and the reversal percentage are only increased slightly. These data are consistent with the idea that pinnae cues, rather than binaural cues, are important for determining source elevation and resolving front-back confusions. The fixed amplitude condition is almost identical to the baseline condition, demonstrating that sound level had virtually no effect on the perception of the direction of the source.

In the low-pass condition, both the standard deviation in elevation and the percentage of front-back reversals are substantially larger than in any other condition. In addition, there is a strong negative bias in elevation, the only major bias seen in the data. These factors contribute to the large angular error found in this condition. Note, however, that after correcting for reversals the standard deviation in azimuth is relatively low. These data are consistent with the idea that high-frequency pinnae cues are necessary for determining source elevation and resolving front-back confusions. Performance in the high-pass condition is only slightly worse than in the baseline condition. The percentage of front-back reversals, however, is increased.

In general, these data are consistent with previous findings about directional localization in the far-field. The summary statistics do not indicate that directional localization is dramatically different in the near-field than in the far-field.

The distance performance results are shown in Figure 5. In this plot, the data for each subject were first sorted by source azimuth and placed into 19 overlapping bins, each containing 10% of the total number of trials. The correlation coefficient between the stimulus distance and the response distance (in logarithmic units) was calculated for each bin. The figure shows the mean correlation coefficient as a function of the mean azimuth value for each bin.

The data show that distance estimation performance depends strongly on azimuth for all conditions. In each case, the correlation coefficient is greater for sources off to the side of the listener than directly in front or behind. For the baseline condition, the correlation coefficient ranges is 0.85 at θ=90°, and only 0.3 directly in front of the subject (near θ=90°). Overall, performance is far worse in the monaural condition than in any other, and is poor in the high-pass condition. The low-pass and baseline conditions are almost identical. For lateral sources, in the fixed amplitude case, the correlation is only slightly better than in the baseline condition, but performance in the fixed amplitude condition is superior to the baseline condition for sources near θ=0°.

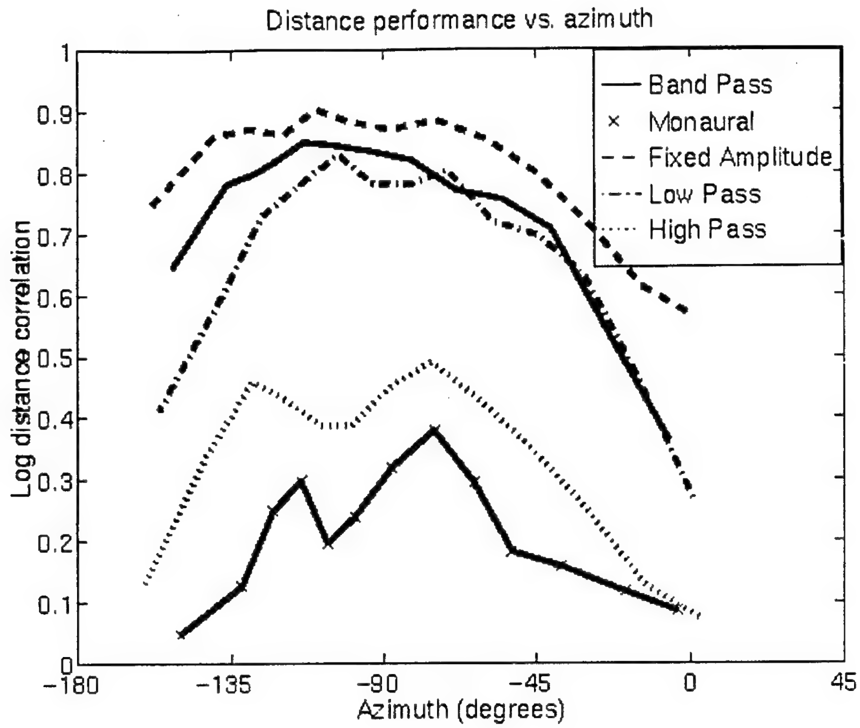

Figure 5: Correlation between Log Stimulus Distance and Log Response

Several conclusions can be drawn from these results. First, that distance perception is robust for lateral sources in the near-field. The correlation coefficient was approximately 0.85 for sources off to the listener's side and at these locations the responses were found to be essentially unbiased, indicating reasonably accurate distance estimation by the subjects. This performance is far better than the distance accuracy reported in far-field distance experiments with random-amplitude stimuli. Second, binaural cues appear to be important for near-field distance perception, which would explain the extremely poor distance performance in the monaural case, as well as the poor performance near $\theta=0°$ where binaural cues are weakest. Third, low low-frequency IID's appear to be an important factor in near-field distance perception. Low-pass filtering of the signal below 3 kHz did not significantly impede distance perception, but high-pass filtering substantially decreased performance. Finally, it is interesting to note that distance accuracy was not well correlated with directional accuracy. Angular error in the low-pass condition was much worse than in the baseline condition, but distance performance is nearly identical. Angular performance in the high-pass condition was only slightly worse than in the baseline condition, but distance perception was extremely poor.

Mathematical Modeling of Near-Field Transfer Functions

In order to better understand the cues available for near-field localization, we have developed a model of near-field HRTFs with the assumption of a spherical head. Rabinowitz et. al. (1993) described a model (based on derivations by Morse and Ingard, 1968) to calculate the pressure generated on the surface of a sphere by a velocity point source at arbitrary distances from the sphere. This model was originally developed to examine the relationship between frequency-scaled HRTFs and actual HRTFs for a magnified head, but it is equally applicable for modeling HRTFs for nearby point sources. According to this model, the pressure on the surface of a sphere due to a nearby point source with volume velocity $u_0 e^{j2\pi ft}$ is

$$\mathbf{P}_s(r, a, \theta, f) = \frac{\rho_0 c u_0}{2\pi a^2} \sum_{m=0}^{\infty} (m + \frac{1}{2}) L_m(\cos\theta) \frac{H_m(kr)}{H'_m(ka)} e^{j2\pi f t},$$

where r is the source distance (to the center of the sphere), a is the radius of the sphere (9 cm), θ is the angle between the point on the surface of the sphere and the direct path to the source, f is the frequency, and k is the wave number $2\pi f/c$. The constant ρ_0 is the density of air (1.18 kg/m^3), and c is the velocity of sound (343 m/s). $L_m(\cos\theta)$ is the Legendre polynomial function in $\cos\theta$, $H_m(kr)$ is the spherical Hankel function in kr, and H_m' is the derivative of the spherical Hankel function with respect to ka.

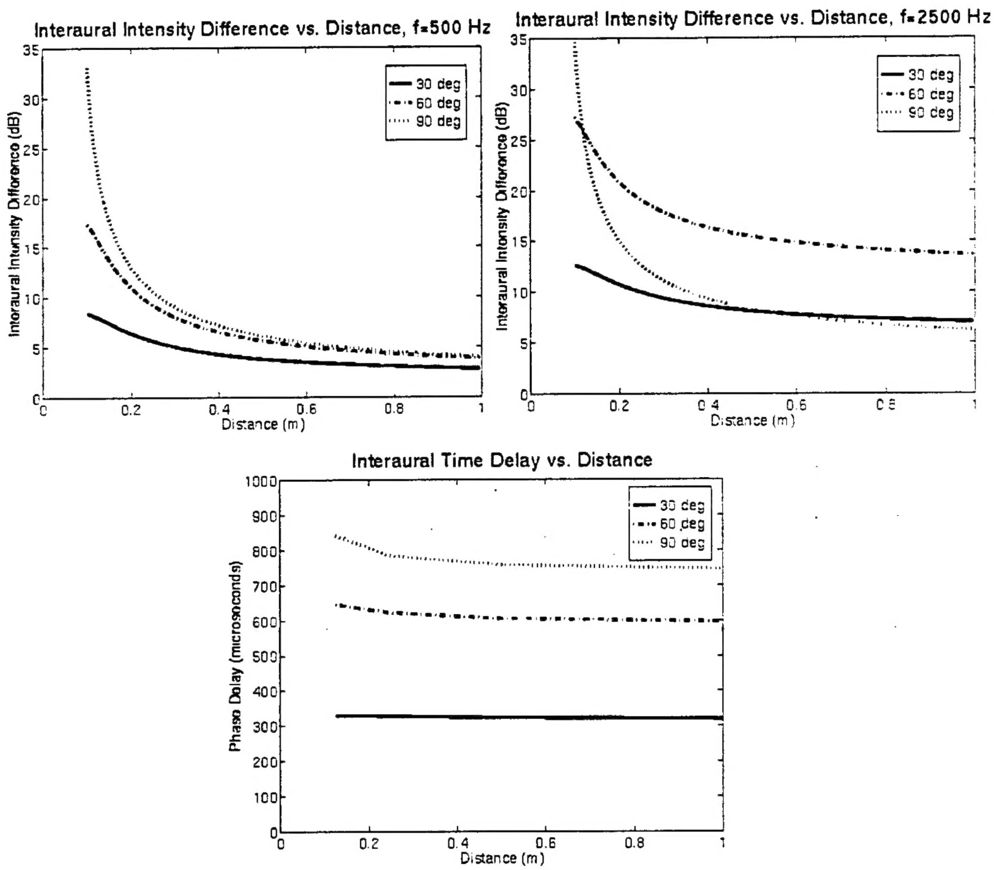

Figure 6: Interaural intensity difference and interaural time delay vs. distance

This equation can be used to determine the interaural intensity differences (IIDs) and interaural time delays (ITDs) for any direction and any distance relative to the head simply by taking the ratio of the pressure at the left ear to the pressure at the right ear. Figure 6 summarizes the important changes in perceptual localization cues that occur in the near field. This figure shows IID and ITD as a function of source distance at three different locations in azimuth. The IIDs increase rapidly as distance decreases, especially at distances less than 0.5 m, while the ITDs increase only slightly. The differences between the ITD and IID appear even more dramatic, however, when perceptual sensitivity is considered. Hershkowitz and Durlach (1969) found that listeners could discriminate changes in IID on the order of 0.8 dB over a broad range of IIDs, indicating that the changes in IID from 0.125 m to 1 m may be as large as 30 JNDs. Conversely, The JND for ITD was approximately 15 μs at ITDs below 400 μs, but increases rapidly for ITDs greater than 400 μs.

Near-field ITDs are primarily dependent on distance near $|\theta|=90°$ where the ITDs are greater than 700 μs and sensitivity to changes in ITD is low. Thus the changes in ITD with distance span, at most, a few JNDs. This analysis indicates that listeners will perceive changes in the IID with distance in the near-field, but will be insensitive to changes in ITD.

Summary

By comparing results across subject populations, we have successfully demonstrated that across-day learning occurs in adaptation paradigms. This finding is important, in that it has direct implications for the utility of training subjects in adaptation paradigms, and the extent to which such training persists.

In order to perform near-field localization studies, we developed and refined a response method that allows quick and accurate measures of perceived location of sound sources. This method was then used to test localization performance in a number of conditions. Results indicate that there are robust distance cues for sources relatively close to the listener which are derived from binaural, low-frequency information. These psychophysical results are supported by analysis of the acoustic distance cues that arise in the near field, which show significant interaural level differences that vary systematically with distance occur in the near field.

References

Coleman, P.D. (1963) An analysis of cues to auditory depth perception in free space. Psychological Bulletin, 60, 302-315.

Hartley, R.V.L. and Frey, T.C. (1921) The binaural location of pure tones. Physics Review, 18, 431-442.

Hershkowitz, R.M. and Durlach, N.I. (1969) Interaural time and amplitude JNDs for a 500-hz tone. Journal of the Acoustical Society of America, 46, 1464-1467.

Makous, J.C., and Middlebrooks, J.C. (1990) Two-dimensional sound localization by human listeners. Journal of the Acoustical Society of America, 87, 2188-2200.

Morse, P.M. and Ingard, K.U. (1968) Theoretical Acoustics. McGraw-Hill, 1968.

Stuart, G.W. (1911) The acoustic shadow of a rigid sphere with certain applications in architectural acoustics and audition. Physics Review, 33, 467-479.

Wightman, F.L., and Kistler, D.J. (1989) Headphone simulation of free-field listening II: Psychological validation. Journal of the Acoustical Society of America, 85, 868-879.

Personnel Supported

The personnel supported by and/or associated with this project are Principal Investigators Nat Durlach, Dick Held, and Barbara Shinn-Cunningham; graduate students Doug Brungart, Salim Kassem, Greg Lin, John Park, Heather Hollett, and Elias Vyzas; and undergraduate Michael DeSenna.

Publications

Shinn-Cunningham, B. G., Durlach, N.I., and Held, R. M. (1998). "Adapting to supernormal auditory localization cues I: Bias and resolution," J Acoust Soc Am, **103**, 3656-3666.

Shinn-Cunningham, B. G., Durlach, N.I., and Held, R. M. (1998). "Adapting to supernormal auditory localization cues II: Constraints on adaptation of mean response," J Acoust Soc Am, **103**, 3667-3676.

Brungart, D. S. and W. M. Rabinowitz (1999). "Auditory localization of nearby sources I: Head-related transfer functions." Journal of the Acoustical Society of America **106**(3): 1465-1479.

Brungart, D. S. and N. I. Durlach (1999). "Auditory localization of nearby sources II: Localization of a broadband source in the near field." Journal of the Acoustical Society of America **106**(4): 1956-1968.

Brungart, D. S. (1999). "Auditory localization of nearby sources III: Stimulus effects." Journal of the Acoustical Society of America **106**(6): 3589-3602.

Shinn-Cunningham, B. G. (2000). "Adapting to remapped auditory localization cues: A decision-theory model." Perception and Psychophysics, **62**(1), 33-47.

Interactions/Transitions

a. Participation/presentations at meetings, conferences, seminars, etc.

Brungart, D. S. and W. R. Rabinowitz (1996). Auditory localization in the near field. Third International Conference on Auditory Display, Palo Alto, CA.

Shinn-Cunningham, B. G., Zurek, P. M., Stutman, E. R., and Berkovitz, R. (1996). "Perception of azimuth for sources simulated using two loudspeakers in natural listening environments", Mid-Winter Meet Assoc Res Otolaryng, 4-8 February 1996.

Shinn-Cunningham, B. G. (1996). "Modeling dynamic changes in auditory localization," talk presented at the Department of Cognitive and Neural Systems Colloquium, Boston University, Boston, MA, 8 May 1996.

Shinn-Cunningham, B. G. (1996). "A dynamic, psychophysical model of adaptation in auditory localization experiments," talk presented at the Department of Psychology Colloquium, Brandeis University, Waltham, MA, 12 September 1996.

Shinn-Cunningham, B. G. (1996). "Modeling resolution in auditory localization experiments," talk presented at the Department of Biomedical Engineering Colloquium, Boston University, Boston, MA, 25 September 1996.

Shinn-Cunningham, B. G. (1996). "Designing effective virtual environments," Proc 1996 Assoc Speech Hear Res, Seattle, WA, 21-24 November 1996.

Brungart, D.S., and Rabinowitz, W.M. (1996) Auditory Localization in the Near-Field. Proceedings of the Third International Conference on Auditory Display, Ed. S. Frysinger, Santa Fe: Santa Fe Institute, 123-126.

Brungart, D.S., Rabinowitz, W.M. and Durlach, N.I. (1996) Auditory localization of a nearby point source. Journal of the Acoustical Society of America, 100(4), 2593.

Shinn-Cunningham, B. G. (1997). "Adapting to discrepant information in multimedia displays." J Acoust Soc Am, **102**, 2237, San Diego, CA, 1 – 5 November 1997.

Shinn-Cunningham, B. G., Zurek, P. M. (1997). "Development of an audiometric localization test," Proc 1997 Assoc Speech Hear Res, Boston, MA, 20-23 November 1997, 231.

Shinn-Cunningham, B. G., Kulkarni, A. (1997). "Localization of median-plane sources: Apparent location of simultaneous sources in anechoic space", Mid-Winter meeting of the Association for Research in Otolaryngology, 2-7 February 1997, 207.

Shinn-Cunningham, B. G., & Lin, G. (1997). Learning new stimulus-response relationships in a localization paradigm: Modeling bias and resolution. Journal of the Acoustical Society of America, 101(5), 3193.

Shinn-Cunningham, B. G., & Park, J. (1997). Changes in auditory localization responses are mediated by visual input. Journal of the Acoustical Society of America, 101(5), 3193.

Zurek, P. M., & Shinn-Cunningham, B. G. (1997). Analysis of phantom source localization cues. Journal of the Acoustical Society of America, 101(5), 3085.

Shinn-Cunningham, B. G. (1997) "Sensorimotor adaptation in auditory localization paradigms: Constraints on learning," talk presented at the Department of Biomedical Engineering Colloquium, Boston University, Boston, MA, 11 July 1997.

Shinn-Cunningham, B. G. (1998). "Binaural models and the precedence effect," Mid-Winter Meet Assoc Res Otolaryng, St. Petersburg Beach, FL, 15-19 February 1998.

Shinn-Cunningham, B. G. (1998). "How experience affects resolution and bias in auditory localization," talk presented at the Eaton-Peabody Laboratory Seminar, May 9, 1998.

Shinn-Cunningham, B. G. (1998). "Plasticity in spatial auditory perception," talk presented at the Boston University Brain, Behavior, and Cognition Colloquium, 3 April, 1998.

Brungart, D. S. (1998). Three-dimensional auditory localization of nearby sources. 42nd Annual Meeting of the Human Factors and Ergonomics Society.

Dizio, P., Lackner, J. R., Held, R., Shinn-Cunningham, B. G., and Durlach, N. I. (1998). "Head-centric auditory localization is affected by gravitoinertial force magnitude and direction," 20th Ann Meet Soc Neurosci, Los Angeles, CA, November 7-12.

Brungart, D. S. (1998). "Control of perceived distance in virtual audio displays." Proceedings of the 20th International Conference of the IEEE Engineering in Medicine and Biology Society **20**(3/6): 1101-1104.

Brungart, D. S. (1998). Preliminary model of auditory distance perception for nearby sources. NATO ASI on Auditory Computational Hearing, Il Ciocco, Italy.

b. Consultative and advisory functions to other laboratories and agencies

none

c. Transitions

none

New Discoveries, Inventions, or Patent Disclosures

Brungart, D.S. and Rabinowitz, W.M. Acoustic Point Source. Patent Application 09/168,339, filed October 5, 1998.

Brungart, D.S. Improved Manikin Positioning for Acoustic Measuring. Patent Application 09/140,063, filed August 24, 1998.

Honors and Awards

none